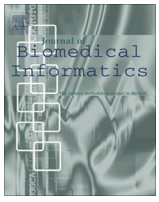
Journal of Biomedical Informatics 48 (2014) 16

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin



Book Reviews

A Gentle Introduction to Support Vector Machines in Biomedicine, Alexander Statnikov, Constantin F. Aliferis, Douglas P. Hardin, Isabelle Guyon. World Scientific (2013). Vols. 1 (200 p.) and 2 (212 p.)

The proliferation of massive amounts of electronic clinical and biomedical data offers unparalleled opportunities for new discoveries with the potential to revolutionize many aspects of biomedical research and patient care. Unlocking this knowledge, however, is only possible through the application of sophisticated data analytics techniques. Such techniques have been developed in the data mining and machine-learning communities, but a significant gap remains between these cutting-edge machine-learning techniques and the methods currently used by mainstream biomedical and clinical researchers. The aim of this book is to bring these two worlds closer together with a conceptually clear and operationally useful description of one of the most popular machine-learning tools: support vector machines.

The book is split into two volumes. The first volume presents the theoretical and conceptual foundations of support vector machines. Building on this foundation, it advances to present methods for classification, regression and clustering. The second volume provides case studies and fully developed applications using the support vector machine methodology, as well as benchmarking results, where support vector machines have been compared to other machine-learning technologies solving real problems.

The first (theory) part of the volume reads like an excellent textbook. It starts by describing convex optimization in a highly accessible manner, exposing the reader only to the concepts necessary to understand the later chapters. The authors then proceed to develop the fundamental concept of support vector machines. While the theory of support vector machines is complex, the authors have managed to strip this theory to its core and provide the clearest description of the material these reviewers have seen. Based on the theory, the authors then develop the most important support vector machine-based techniques for classification, clustering, and regression. While this volume is a collaboration among many well-regarded authors, it reads very clearly and consistently: the presentation and writing flows uniformly across the chapters and they logically build upon one another. This volume is an outstanding introduction to support vector machines and, unlike many other texts, it truly lives up to its title: it is a "gentle" yet comprehensive introduction.

While the first volume reads like a textbook, the second volume has the feel of an edited book by multiple authors. Indeed, the chapters in this volume are contributed by numerous scientists. The greatest strength of the second volume is its first chapter, which provides an overview of the support vector machine methodology (superfluous if you have read the first volume), evaluation methodologies, and competing machine-learning techniques. Then it proceeds to describe individual case studies by different authors, with examples drawn almost exclusively from the bioinformatics domain. These sections read like individual scientific papers, lacking the consistency in writing style and the logical progression that we came to expect having read the first volume.

In summary, the first volume is a most outstanding "gentle" introduction to support vector machines, written by authors with excellent credibility who have substantially contributed to the development of support vector machine-based techniques. The first volume should be recommended to scientists who seek to learn cutting-edge machine-learning techniques or to extend their repertoire of analysis techniques that work well in practice. The second volume also contains useful information, including variable selection techniques and details of complete applications of the support vector machine classifier. However, this volume has a difficult time living up to the very high standard of consistent and clear writing that the first volume has set.

Gyorgy Simon
Institute for Health Informatics, University of Minnesota, Minneapolis, MN, United States
E-mail address: simo0342@umn.edu

Genevieve B. Melton
Institute for Health Informatics, Department of Surgery, University of Minnesota, Minneapolis, MN, United States
E-mail address: gmelton@umn.edu

Available online 17 March 2014

http://dx.doi.org/10.1016/j.jbi.2014.02.012